

го искусства. Вместе с тем Российская государственная цирковая компания смогла использовать медиасреду в целях привлечения аудитории. Этот опыт можно использовать и в дальнейшем, например, для знакомства с теорией и историей цирка или проектов, не требующих непосредственного присутствия публики, но для полноценной и эффективной работы и для самого существования циркового искусства все-таки необходимы «живые» выступления и реальные зрители.

Библиографические ссылки

1. Советский цирк на пяти континентах / Ф. Бардиан. М. : Искусство, 1977. С. 6.
2. Новости // ФКП «Росгосцирк» : [офиц. сайт]. URL: http://www.circus.ru/press-service/news/24239/?sphrase_id=34465 (дата обращения: 22.10.2020).

УДК 81.322

М. Д. Маслова

Уральский федеральный университет, Екатеринбург

Корпус диалектных текстов как возможность удаленной работы с материалом: особенности морфологической разметки

В данном исследовании создание корпуса диалектных текстов рассматривается как шаг к тому, чтобы сделать работу с фольклорными материалами возможной вне зависимости от местонахождения исследователя, что особенно актуально в период карантина. Освещаются проблемы морфологической разметки диалектных текстов и способы их решения.

Ключевые слова: корпусная лингвистика, диалектный корпус, морфологическая разметка, морфологический анализатор *Mystem*

Dialect corpus as a possibility of remote materials' studying: the peculiarities of morphological markup

In this paper dialect corpus is considered as the way to make researcher's work independent of his working place what is a topical matter in our quarantine time. The peculiarities of dialect corpus's morphological markup are outlined in this work.

Keywords: corpus linguistics, dialect corpus, morphological markup, morphological analyzer *Mystem*

Работа с языковым материалом является неотъемлемой частью исследовательской деятельности лингвиста. Обращаясь к проблематике региональных текстов и возможных диалектных особенностей речи, необходимо внимательно изучить соответствующие теме материалы фольклорных экспедиций. Традиционно они хранятся в виде рукописей, рассортированных по папкам в зависимости от места и времени сбора данных. Однако на сегодняшний день у такого способа хранения можно найти немало недостатков. Главным из них является тот факт, что работа с данными в рукописном формате предполагает присутствие исследователя в университете или хотя бы возможность забрать материалы с собой, сделать их фото.

За время пандемии доступ в университет ограничивался несколько раз, что повлекло за собой трансформацию всей системы работы лингвистов с текстовыми материалами (от картотек и словарей до различных собраний текстов). Исследователи в области языкознания были вынуждены обратиться к существующим в электронном виде данным. Безусловно, многие словари, картотеки, аннотированные собрания текстов в настоящее время находятся в открытом доступе в Интернете. Однако этого нельзя сказать о диалектных текстах, большая часть которых существует лишь в рукописном формате. Видя неустойчивость ситуации в мире, мы осознаем, что для непрерывной научной деятельности нужен новый формат представления материала, позволяющий работать с диалектными

текстами многим людям вне зависимости от их местонахождения. В реализации данной идеи полезным будет опыт корпусной лингвистики. При разнообразии подходов к созданию диалектных корпусов наиболее распространенной концепцией является морфологическая. Самым обсуждаемым является диалектный подкорпус НКРЯ и его система диалектных помет, описанная А. Б. Летучим [1, с. 215–232]. Несмотря на критику некоторых исследователей [2, с. 278–283], нам видится целесообразным придерживаться концепции данного корпуса. Работая в таком русле, исследователи УрФУ со временем могут представить аннотированное собрание уральских текстов для пополнения самого НКРЯ.

Безусловно, работа по разметке большого количества текстов невозможна, пока не будет создан пилотный проект. В качестве материала для создаваемого корпуса были отобраны тексты, сохранение которых имеет особую важность. Выбранные нами материалы были собраны в 1960–1970-е гг. в 13 близко расположенных деревнях, которые на сегодняшний день находятся на грани исчезновения. Таким образом, введение этих текстов в научный оборот также позволит сохранить культурное наследие уходящих в прошлое русских поселений.

На этапе первичной обработки материалов по вышеуказанному критерию были отобраны 118 текстов девяти жанров. На основе информации, указанной на бумажных носителях, были собраны данные для метаразметки корпуса: время и место сбора материала; имя, возраст, год рождения информанта; имя и фамилия собирателя; жанр текста.

Однако главный интерес представляло создание морфологической разметки, причем первичное аннотирование было осуществлено при помощи морфоанализатора *Mystem*. Несмотря на то, что создатели программы считают возможным формирование морфологических гипотез для незнакомых слов [3], нами был выявлен ряд текстоформ, проанализированных неверно в силу их диалектного характера. Сходство некоторых проблемных случаев друг с другом позволило разделить их на группы, для каждой из которых было выработано решение с опорой на систему диалектных помет НКРЯ.

Первая категория — имена собственные — реалии местности, в которой живет информант: *дозморята, гамаюны*. Отметки о диалектности здесь не требуются. Исследователь должен указать верную начальную форму слова и добавить помету *geo, persn* или *fam* в зависимости от значения слова.

Следующая группа требует не только ручной разметки, но и работы со Словарем русских говоров Среднего Урала. Речь идет о диалектных лексемах, не имеющих даже отдаленной схожести с лексемами литературного языка: *отгаркиваться, тихомольно*. Диалектные лексемы вручную отмечают пометой *diallex*.

Третья группа слов связана с диалектным словообразованием. В подобных случаях к литературному корню присоединяется словообразовательный формант, в целом существующий в литературном языке, но не используемый с данным корнем: *полушалчик, сповенчаться*. Подобные словоформы помечаются как *diallex*.

Следующий тип слов объединен таким признаком, как диалектные изменения в морфемах. Наиболее сложными для *Mystem* оказались формы, отмеченные впоследствии как *dialrefltype* (вариативность показателя возвратности) и *dialcontr* (стяжение форм прилагательных): *прибавилася* и *больша́* соответственно. Некоторые формы, получившие помету *dialflex* (диалектное окончание), вообще не были разобраны *Mystem*: *малыи*.

Последняя категория диалектных изменений проявляется исключительно через сочетаемость рассматриваемых форм с другими (отклонение от литературной нормы в категории рода или падежа). Характеристика таких слов будет неполной, если особенности, проявляющиеся в их сочетаемости, не будут отмечены вручную пометами *dialgend* или *dialcase*.

Морфологическая разметка диалектного корпуса ставит исследователя перед рядом сложностей, которые не могут быть решены лишь при помощи программ автоматического аннотирования. Однако подобная работа важна не только потому, что она позволяет структурировать языковые особенности диалектных текстов, но и потому, что дает начало полноценному цифровому представлению фольклорного материала. Это даст возможность исследова-

телям не прекращать всестороннее изучение региональных текстов, несмотря на непредсказуемость сегодняшней ситуации в мире.

Библиографические ссылки

1. Летуший А. Б. Корпус диалектных текстов: задачи и проблемы // Национальный корпус русского языка: 2003–2005. М. : Индрик, 2005. С. 215–232.
2. Крючкова О. Ю., Гольдин В. Е. Диалектный корпус в составе Национального корпуса русского языка: критический анализ // Язык — текст — дискурс: проблемы интерпретации высказывания в разных коммуникативных сферах. Самара : Универс групп, 2011. С. 278–283.
3. Segalovich I. A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine // Yandex : [website]. URL: <https://cache-ekt04.cdn.yandex.net/download.yandex.ru/company/iseg-las-vegas.pdf> (accessed: 12.12.2019).

УДК 159.9

Е. А. Медведева

Уральский федеральный университет, Екатеринбург

Изучение особенностей восприятия учебного материала в онлайн-обучении студентами различных модальностей

Изменения форм обучения в связи с борьбой с COVID-19 актуализировали проблему оценки эффективности удаленного обучения. Основными психотипическими особенностями являются модальности личности, которые отвечают за процесс усвоения информации. В работе представлены результаты исследования студентов различных модальностей, демонстрирующие универсальные сложности каждого психотипа. Знание особенностей усвоения и имеющихся трудностей в онлайн-обучении позволит индивидуализировать пути и методы работы с учебным материалом.

Ключевые слова: пандемия, модальность, онлайн-обучение